NOAA-CIRES Climate Diagnostics Center

Advancing understanding and predictions of climate variability

Model errors in ensemble forecasts: the structure of errors from unrepresented scales

Tom Hamill and Jeff Whitaker

NOAA-CIRES Climate Diagnostics Center

Boulder, Colorado

Tom.Hamill@noaa.gov

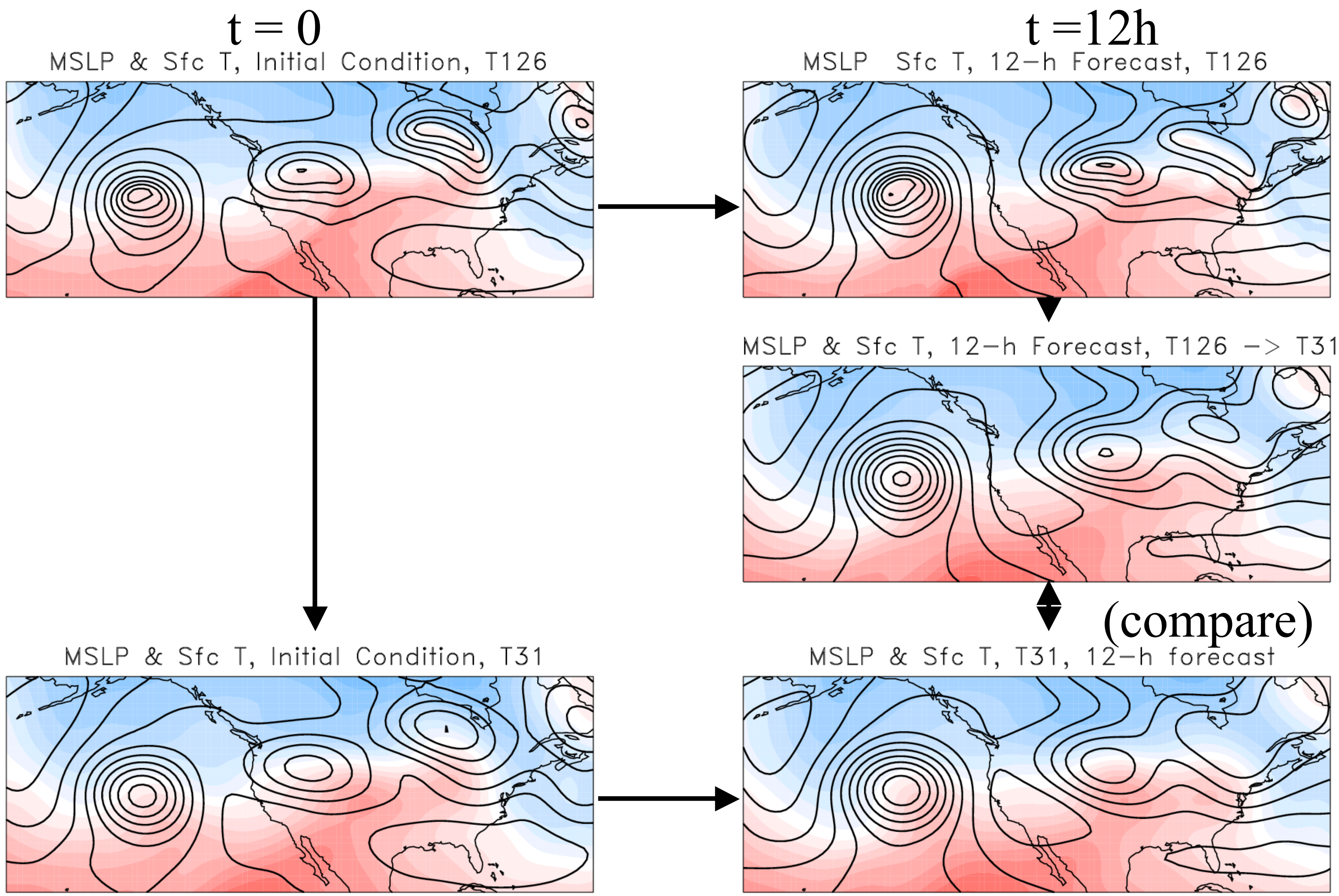
Motivation

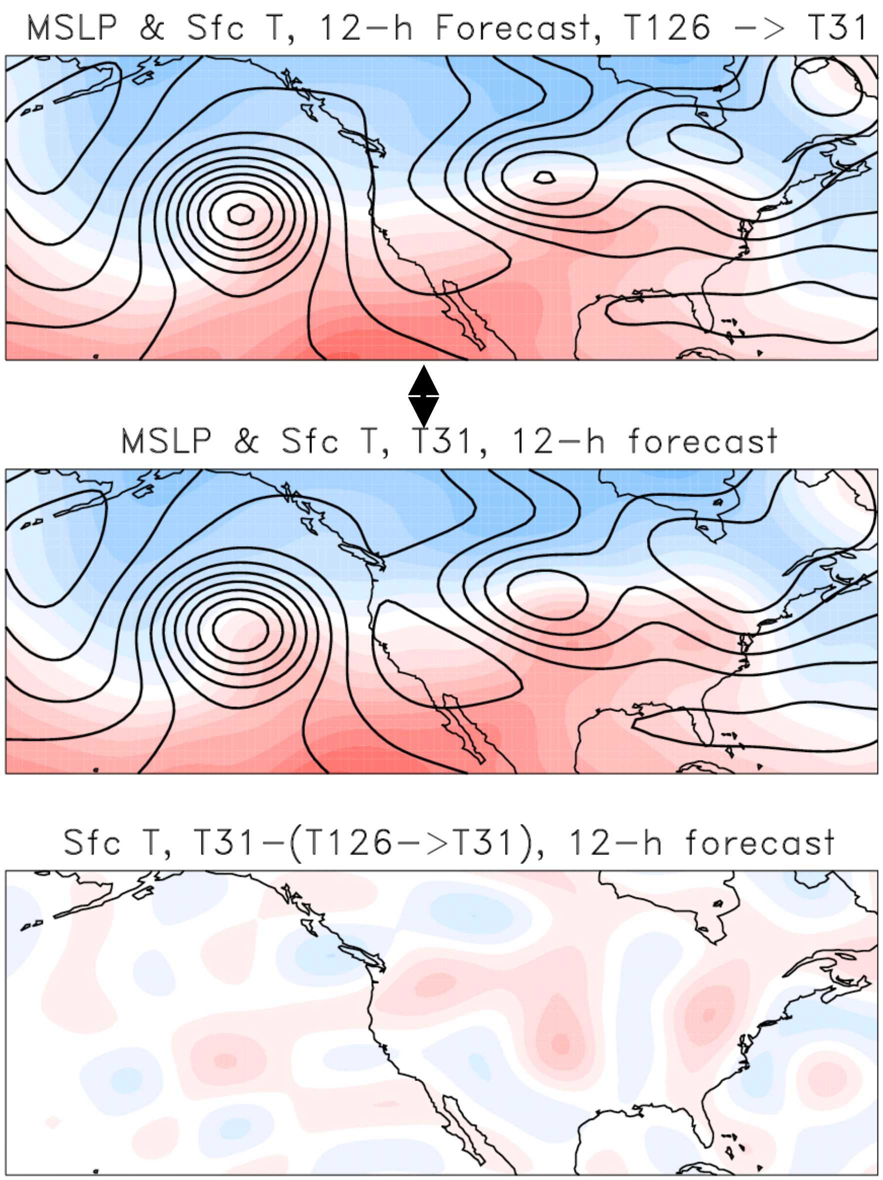
- NWP errors have two sources
  - (1) Growth of initial condition errors (chaos)
  - (2) Model error (insufficient resolution, incorrect parameterizations, etc.)
- Can we understand some general characteristics of model error due to *insufficient resolution*?
- Can we “parameterize” this model error so that ensemble data assimilations are improved?

Experiment Design

- Dry, primitive equation global spectral model, no terrain. Forcing like Held-Suarez (relaxation to zonal temperature profile).
- TRUTH: T126 L30 simulation
- FORECAST: T31 L30 simulation

Evaluating model errors in low-resolution version of high-resolution model

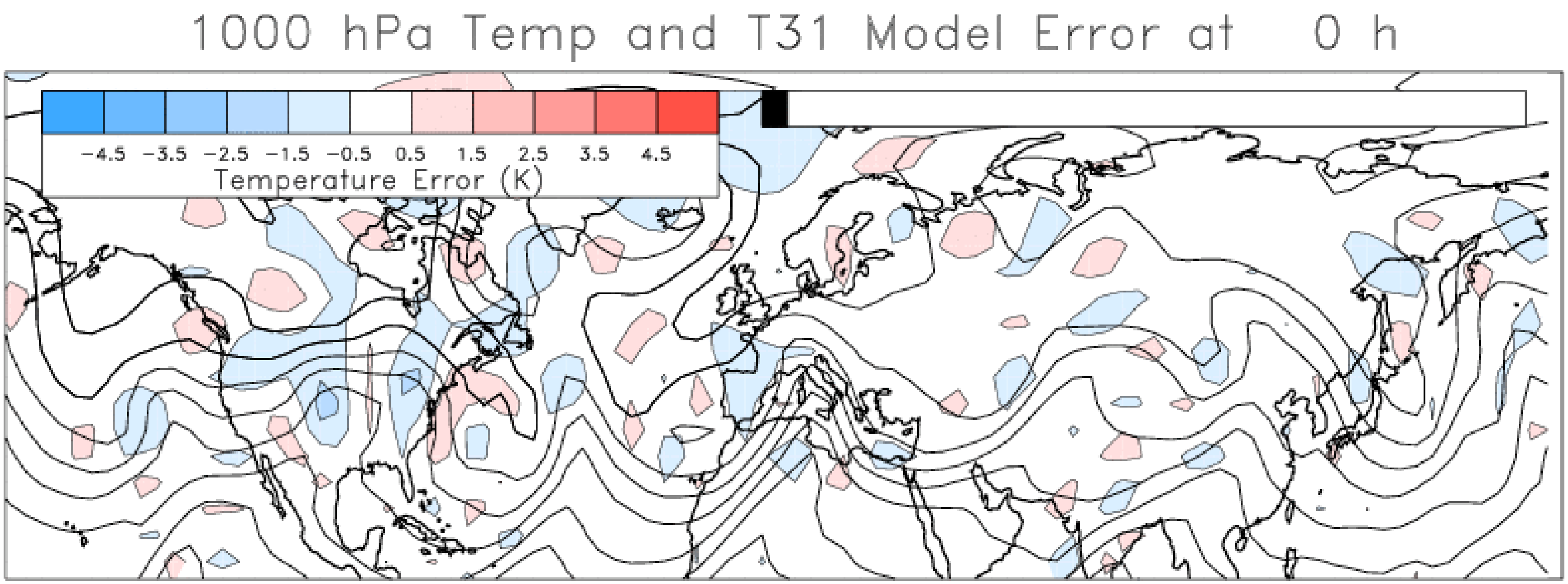




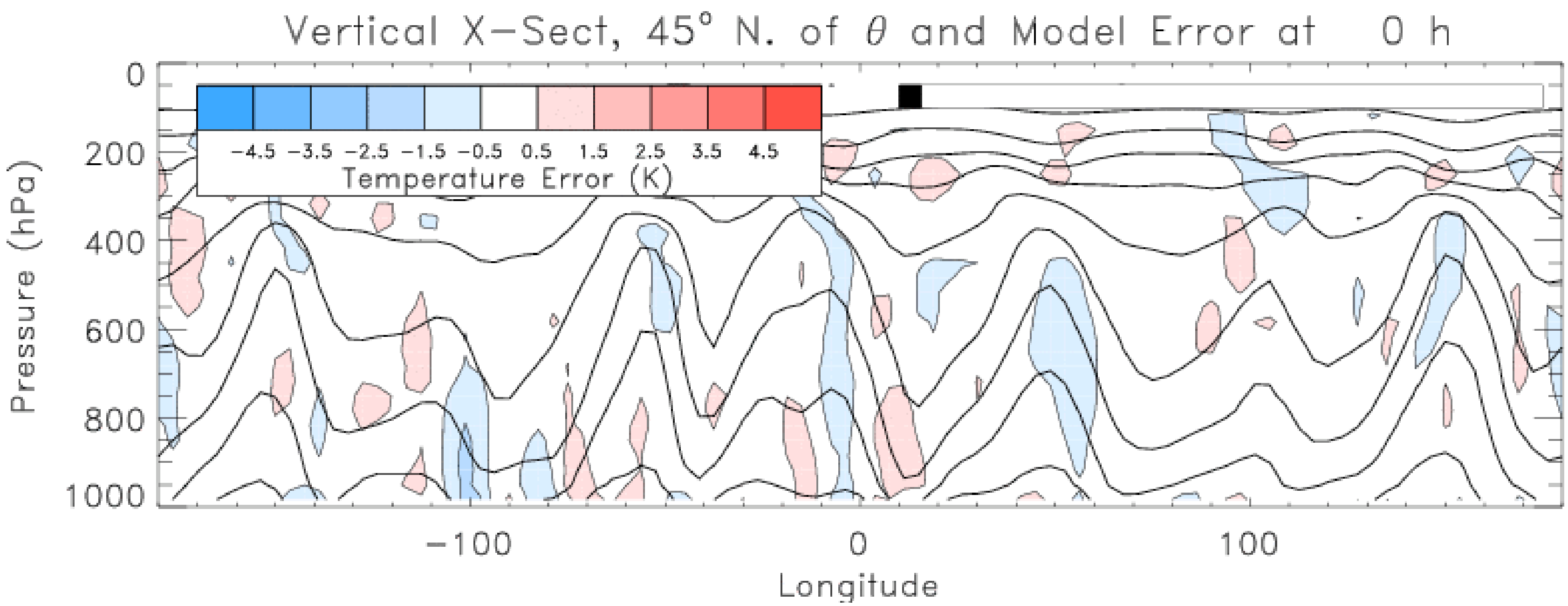
Animate a series of these 12-h forecast errors due to truncation

different model climates?

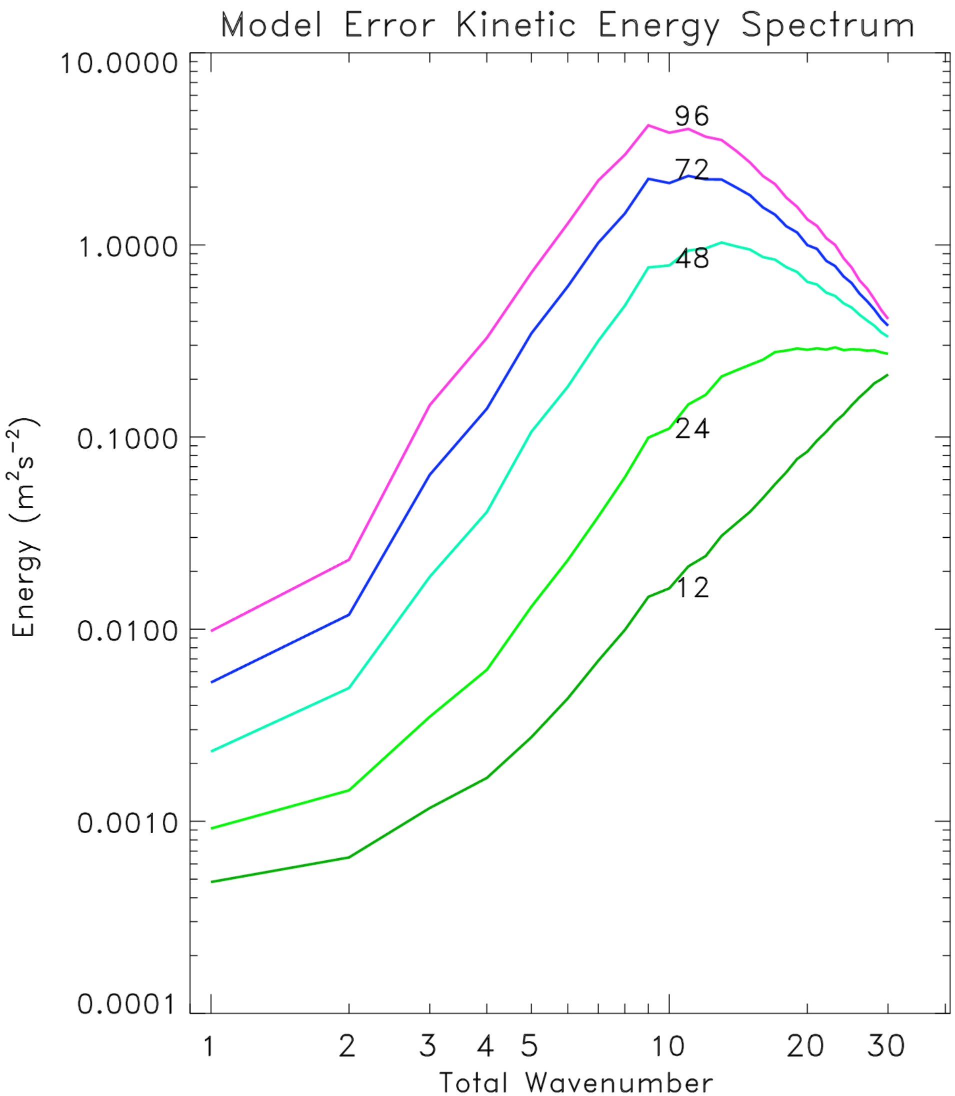
12-h model error due to truncation



12-h model error due to truncation (vertical cross section)

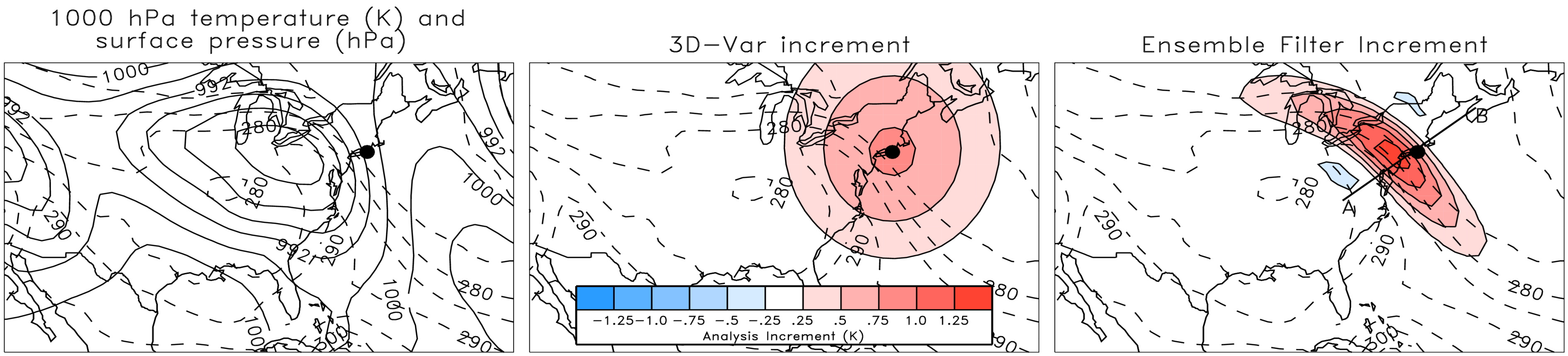


Growth properties of truncation errors



Similar result
in Tribbia and
Baumhefner,
upcoming MWR

Ensemble data assimilation



Specially constructed ensembles of forecasts used to model the forecast-error covariances used in data assimilation. Under the right conditions, (1) an ensemble of perturbed initial conditions will be created that samples the analysis-error covariances, and (2) the ensemble mean analysis will be more accurate than analyzed states from 3D-Var (or perhaps even 4D-Var, in some circumstances).

Data assimilation terminology

- $\mathbf{y}$: Observation vector (raobs, satellite, etc.)
- $\mathbf{x}^b$: Background state vector (1st guess)
- $\mathbf{x}^a$: Analysis state vector"
- H : Operator to convert model state → obs
- $\mathbf{R}$: Observation - error covariance matrix
- $\mathbf{P}^b$: Background - error covariance matrix
- $\mathbf{P}^a$: Analysis - error covariance matrix

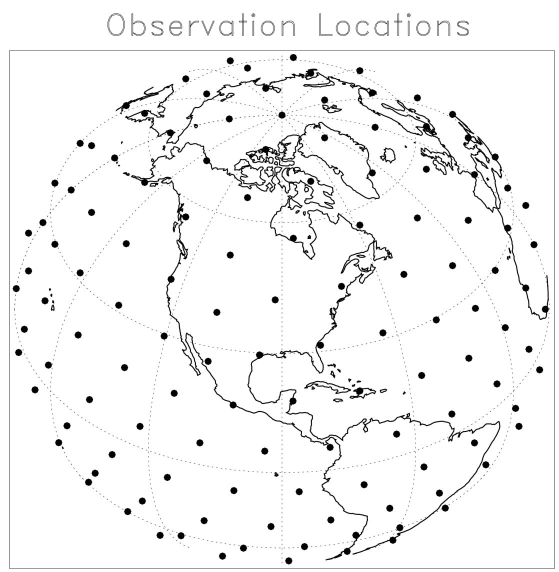
Ensemble Kalman filter equations

$$\mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{K}\Big(\mathbf{y}_i - H(\mathbf{x}_i^b)\Big).$$
$$\mathbf{y}_i = \mathbf{y} + \mathbf{y}_i' \,, \qquad \mathbf{y}_i' \;\sim\; N(0, \mathbf{R})$$
$$\mathbf{K} = \mathbf{P}^b H^{\mathbf{T}}(H\mathbf{P}^b H^{\mathbf{T}} + \mathbf{R})^{-1}$$
$$\mathbf{P}^b = \rho \circ \tfrac{1}{n-1}\mathbf{X}\mathbf{X}^{\mathbf{T}} \,, \qquad \rho \;=\; \text{covariance localiz.}$$
$$\mathbf{X} = (\mathbf{x}_1^b - \overline{\mathbf{x}^b}, \ldots, \mathbf{x}_n^b - \overline{\mathbf{x}^b})$$
$$\mathbf{x}_i^b(t+1) = M(\mathbf{x}_i^b(t)) + e, \qquad e \sim N(0, \mathbf{Q})$$

(We’ll use a slight variant called the “ensemble square-root Filter, or “EnSRF” that doesn’t require perturbed observations)

Ensemble Square-Root Filter (EnSRF) simulations

- 50 members @ T31 L30, 3500 km cov. loc.
- 252 observations of U,V,T at 7 levels, plus SLP; obs = T126 + random error



(obs separated by ~ 1300 km)

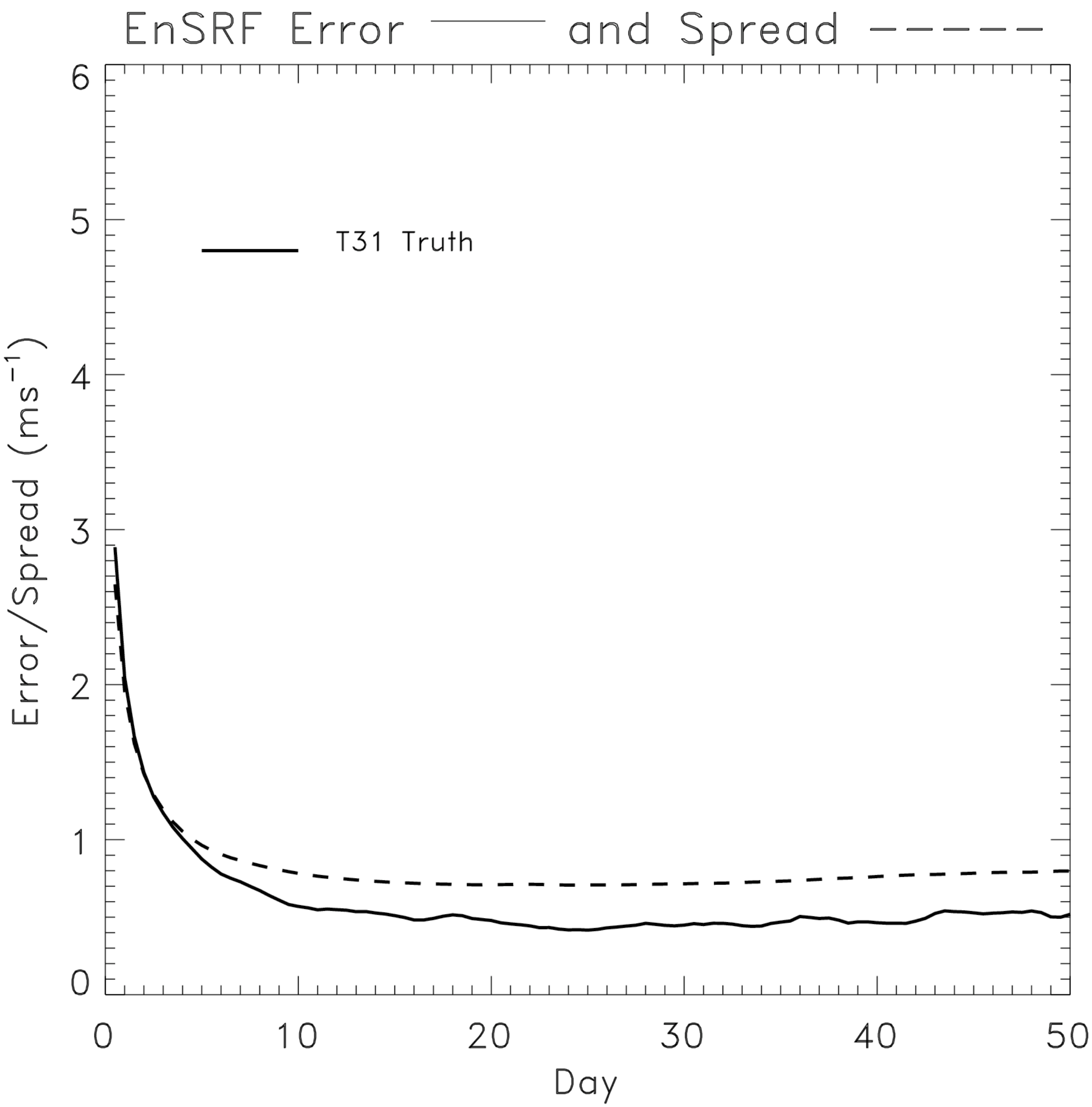
The gamut of simulations

- *Covariance inflation*: $\mathbf{x}_i^b \leftarrow (1+r)\ (\mathbf{x}_i^b - \overline{\mathbf{x}^b}) + \overline{\mathbf{x}^b}, \quad r > 0.0$

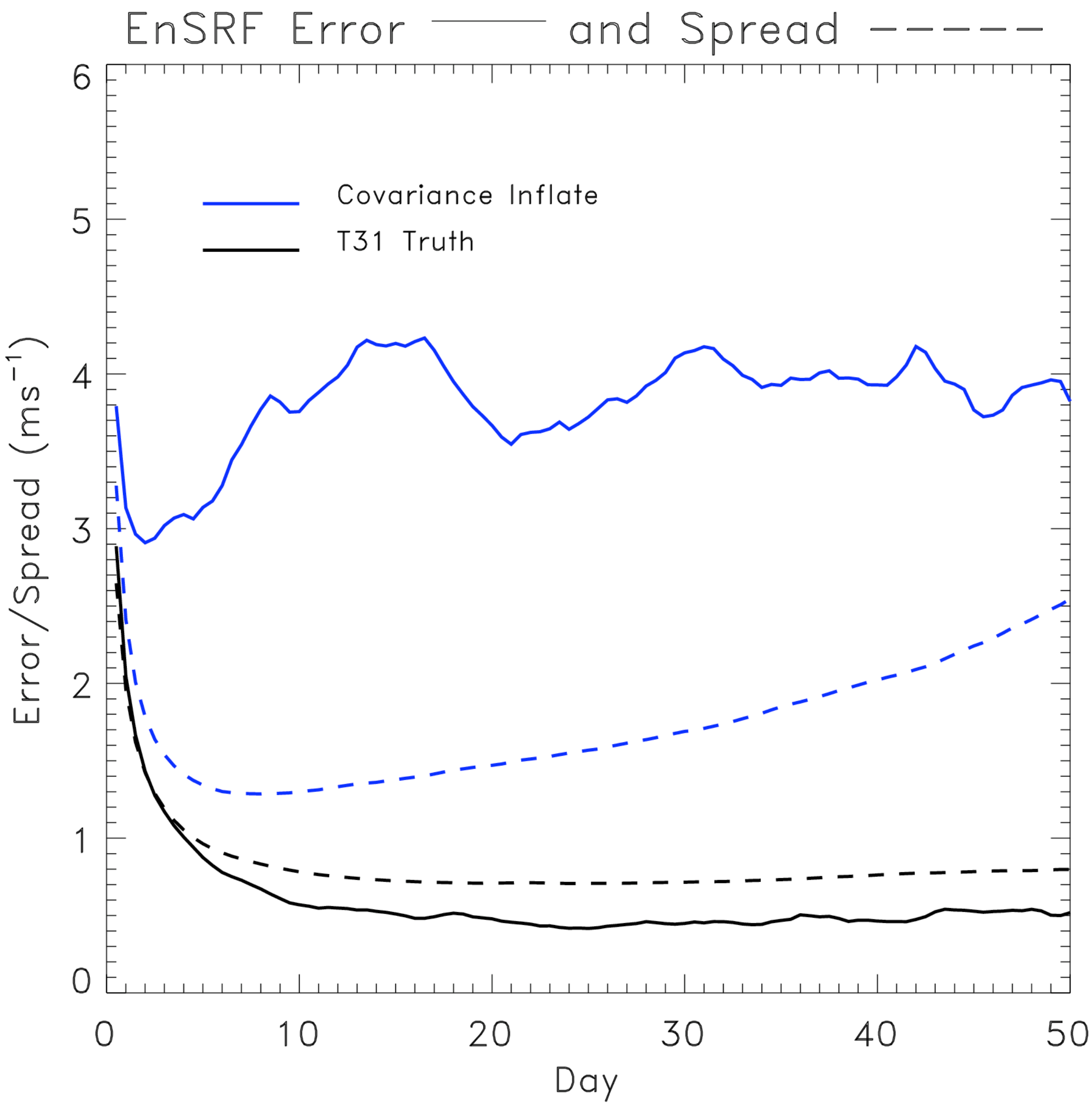
- *Additive error*: $\mathbf{x}_i^b \leftarrow \mathbf{x}_i^b + \mathbf{z}_i\ , \quad \mathbf{z}_i$ is additive model error

- Experiments assimilating T126 obs (model error) :
  - r = 8 % covariance inflation
  - r = 2 % inflation + "unbalanced" additive error
  - r = 2 % inflation + "balanced" additive error

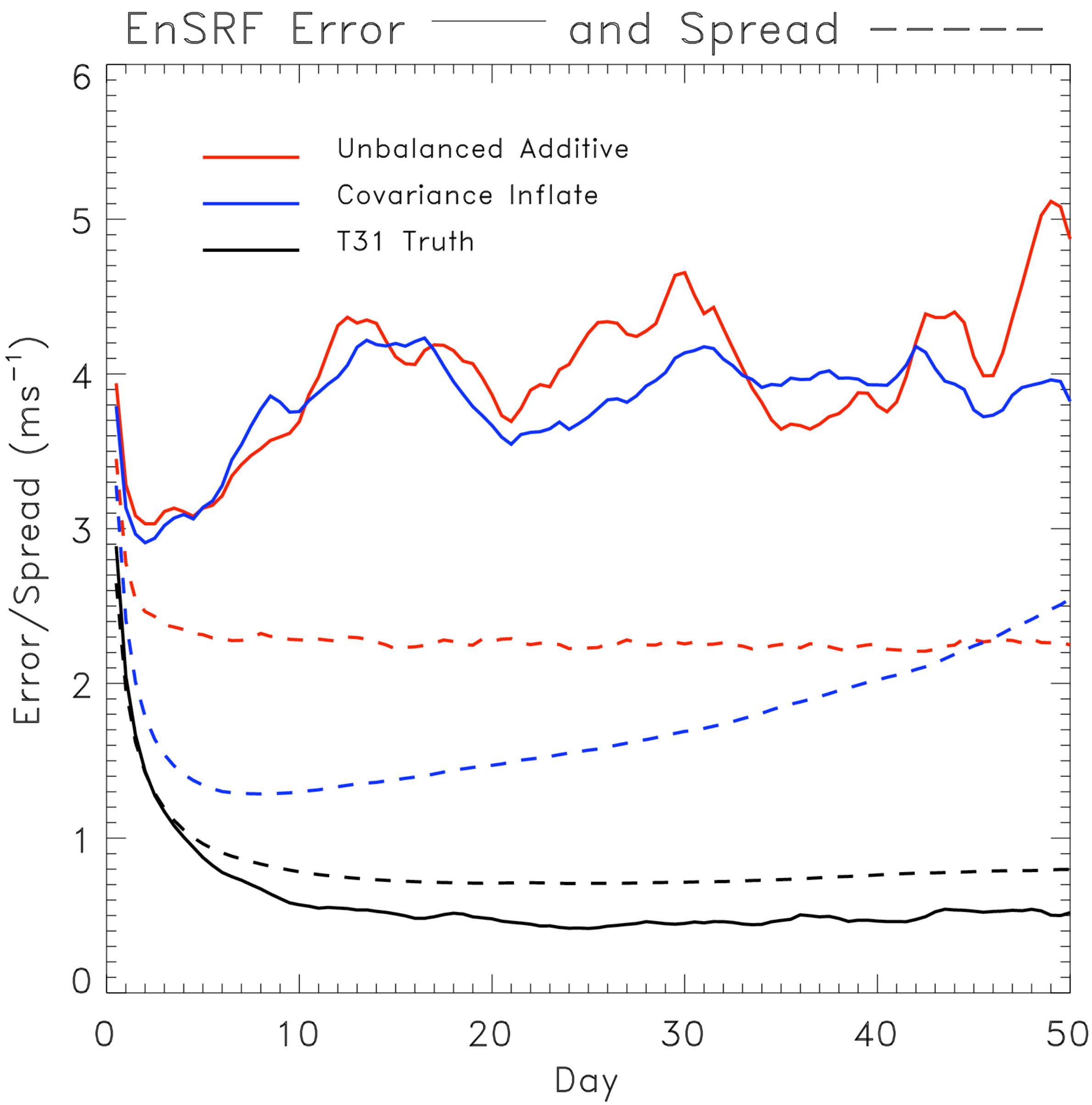
- Experiments assimilating T31 obs (perfect model)
  - r = 4 % inflation

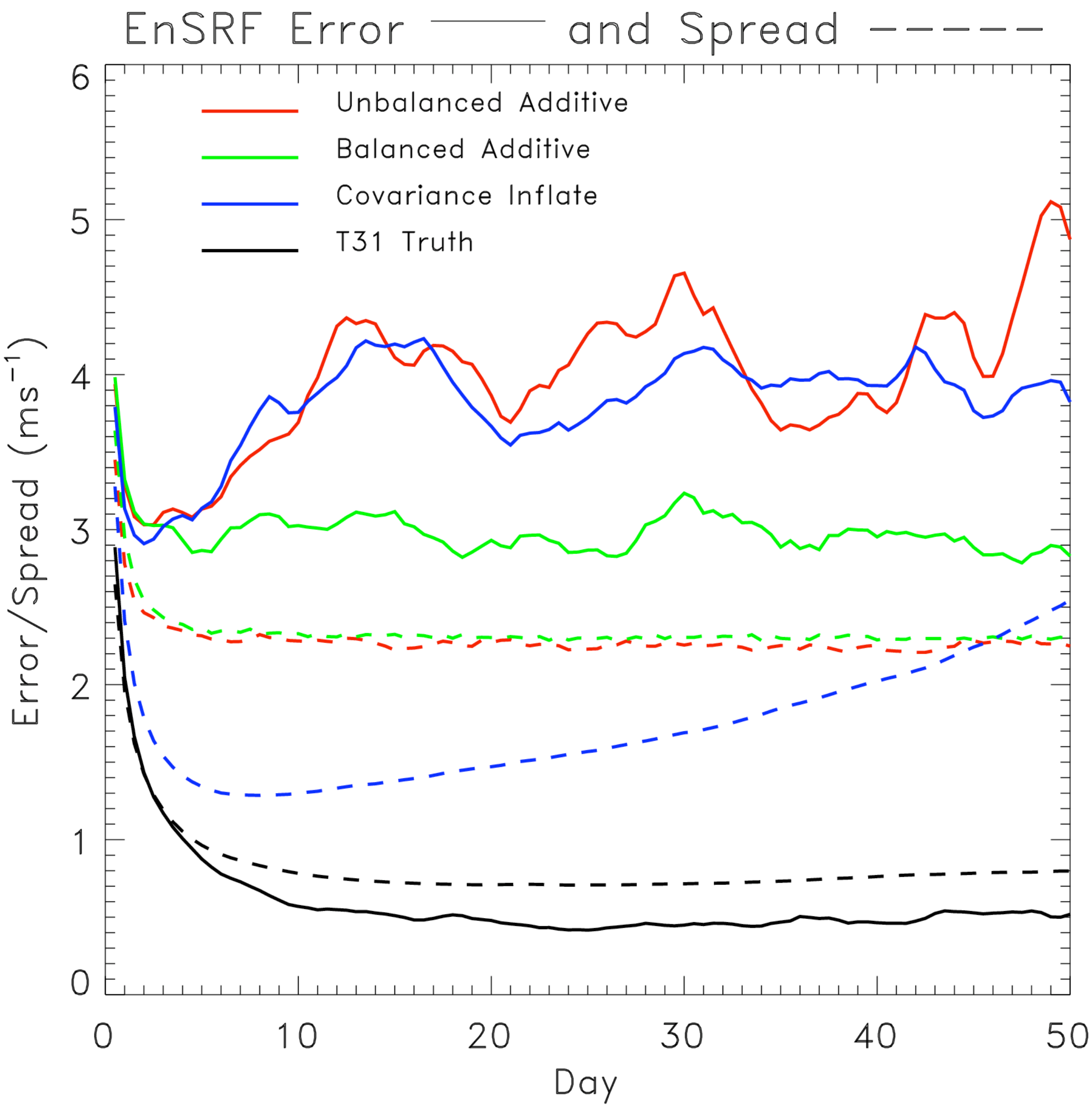


EnSRF Error and Spread
T31 Truth
Error/Spread (ms⁻¹)
Day
0
1
2
3
4
5
6
10
20
30
40
50



EnSRF Error and Spread
Covariance Inflate
T31 Truth
Error/Spread (ms^{-1})
Day
0
10
20
30
40
50
0
1
2
3
4
5
6


EnSRF Error ——— and Spread - - - - -
Unbalanced Additive
Covariance Inflate
T31 Truth
Error/Spread (ms^{-1})
Day
0
1
2
3
4
5
6
10
20
30
40
50

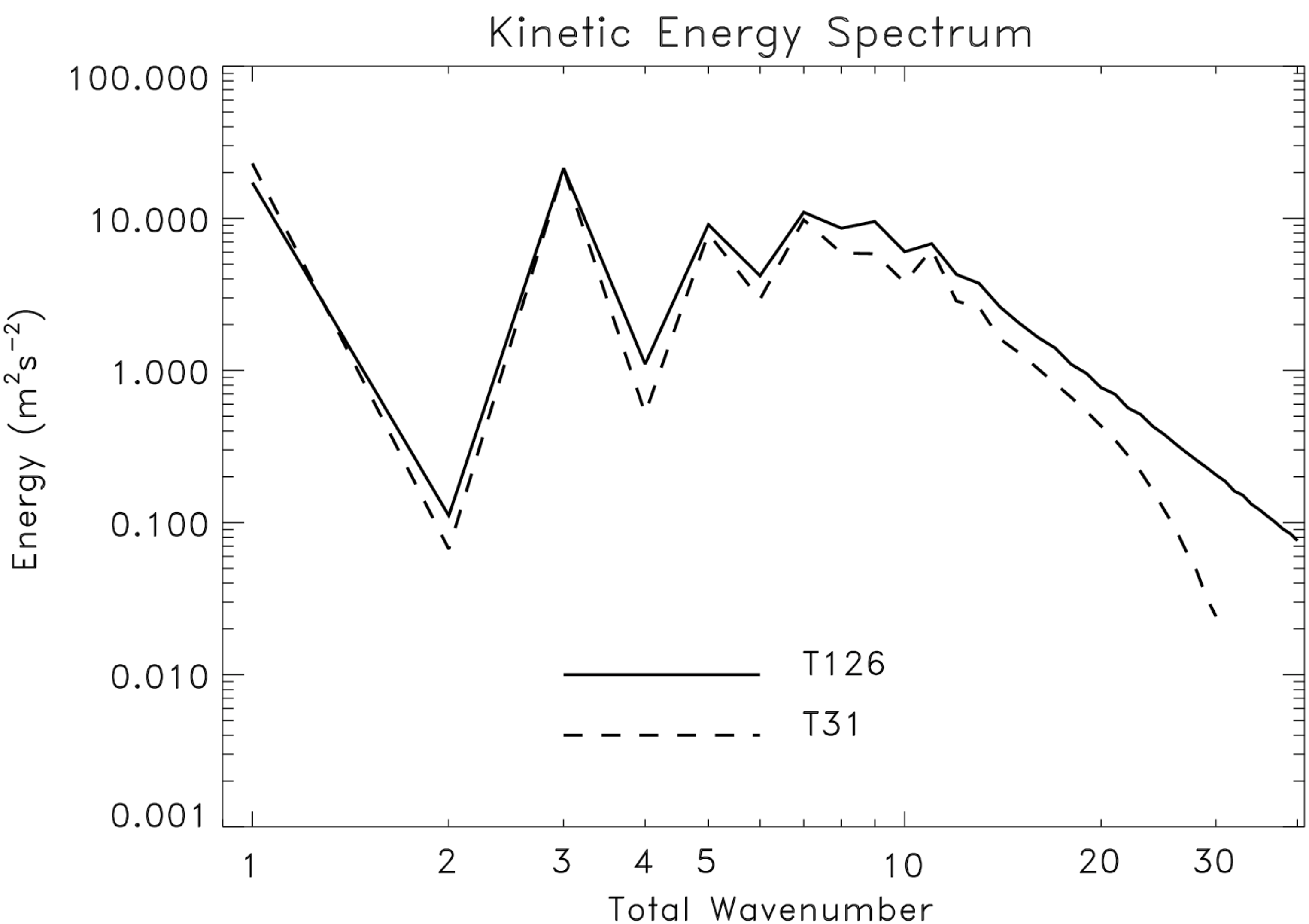

EnSRF Error and Spread
Unbalanced Additive
Balanced Additive
Covariance Inflate
T31 Truth
Error/Spread (ms⁻¹)
Day
0
1
2
3
4
5
6
0
10
20
30
40
50

3D-Var?

Conclusions

- Model errors due to insufficient resolution show some flow dependence and temporal correlation
- Errors start primarily small in scale, grow upscale by 48 h.
- The method of parameterizing model errors in ensemble data assimilations matters; errors can be reduced substantially w. better parameterizations.
- Future: further explorations of model error in more complex models, structure of errors due to parameterization errors.
- http://www.cdc.noaa.gov/~hamill/modelerr.pdf

Some drift of T31 toward different model climate.



Constructing additive model error

- "Unbalanced" additive: see preprint. Sample constructed from a linear combination of singular vectors of model error.
- "Balanced" additive: sample constructed from a random sample of

$T31_{12H}$ - ($T126_{12H}$ --> $T31_{12H}$)

Assimilation experiments with a simpler, cheaper 2-layer PE model